Research Paper | Computer Science | E-ISSN No : 2454-9916 | Volume : 3 | Issue : 3 | Mar 2017

SENTIMENT ANALYSIS ON ONLINE PRODUCT REVIEWS

Sayali Pednekar [1] | Komal Patil [1] | Rutweek Sawant [1] | Trupti Shah [2]

[1] B.E. Student of Computer Engineering, Atharva College of Engineering, Mumbai University, Mumbai, MH, India.

[2] Prof. of Computer Engineering, Atharva College of Engineering, Mumbai University, Mumbai, MH, India.

ABSTRACT

In today's e-commerce world sentiment analysis is salient and the most crucial step because it captures the product users opinion, feelings and belief regarding the respective product specially to determine whether the users attitude is positive, negative, or neutral .This analysis greatly helps the companies to make necessary changes in their product which in return can overcome the flaws that the product is facing and target to a better customer satisfaction.

KEY WORDS: Sentiment analysis; product reviews; opinion mining; feature score.

I. INTRODUCTION

We often observe that the potential customers/ prospects use "product reviews and ratings", very repeatedly. The product reviews are of utmost importance for the buyers to decide based on their concerns regarding product's various aspects for example a phone's screen size, processor speed, memory etc. Hence sentiment analysis of product reviews provides nearly accurate statistics regarding a product, providing an ease to the customers for analyzing the product and zero down his/her search for an online product. The key focus here is efficient feature extraction, polarity classification thereby summarizing positive and negative or neutral polarity.

A core priority in sentiment analysis is to identify the polarity of a given text, sentence, or an entity feature Sentiment analysis is a crucial part of business analytics used by various product manufacturing companies as it helps them to identify satisfaction level of users. Locate the drawbacks in product and find the areas of improvement.

II. LITERATURE SURVEY

The main approaches used to generate extractive summaries are (1) calculating semantics of words(2) lexical chains (3)combinations of heuristics. The traditional text summarization and online product reviews summarization are completely different aspects because during summarizing user reviews, the aim is to gather and identify semantic features of products and later generate a comparative summary of products based on data obtained from feature-wise sentiment classification of the reviews that in turn will help the users and potential buyers in making a decision. It has been demonstrated by the authors that traditional unsupervised text classification techniques like naive Bayes, support vector machine do not perform well on sentiment or opinion classification and hence they pointed out the necessity for feature-oriented classification. Hence the recent research work in opinion mining is mainly oriented towards and summarization and feature based extraction.

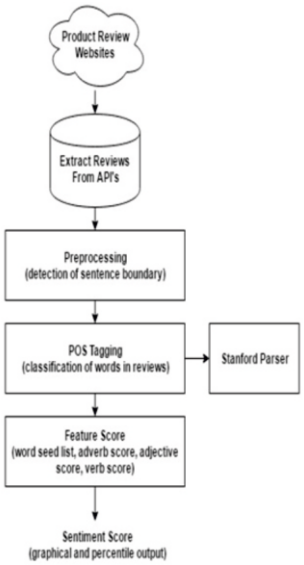

Fig. Architecture

III. PROPOSED SYSTEM

We propose a framework that will summarize as well as classify reviews and extract sentiment score from it. This can be done by crawling websites that provide their API's and categorically post product reviews by actual users. As shown in figure the analysis has four main phases:

- Review's extraction
- Preprocessing
- POS tagging
- Feature score

- **Review's extraction:** Reviews can be pulled of sites which provide their API's openly also some sites do not provide API's openly hence we would require their formal approval and consent and after that we can use them. For a more accurate analysis large number of reviews are required hence the aim should be to pull maximum reviews.

- **Preprocessing:** In preprocessing stage, the acquired reviews are processed in terms of spelling errors as well as sentence boundary detection. It has been observed that many reviewers tend to use slang language and to change the grammatical correct words by adding certain letters to emphasis their feelings e.g. Greatttttt,awwwwesome etc. Hence such kind of words should be processed by clubbing repeating letters into a single letter e.g. Greatttttt→Great.

- **POS tagging:** Reviews contain various features about products such as battery, display, processor etc. These features are tagged as nouns which are considered for identifying features polarity. These features are always described by various adjectives and adverbs which play a crucial role in classifying sentiment.

  In order to extract nouns, adjectives adverbs from reviews POS tagging is performed with the help of a Stanford POS tagger .A Stanford POS tagger is a software that takes text as input after that it assigns each word its respective part of speech such as word, noun, adjective etc.

- **Feature Score:** Once the words are classified, we then use SentiWordNet to score the classified words. SentiWordNet is a lexical resource for opinion mining. SentiWordNet is used to assign synset of WordNet three sentiment scores: positivity, negativity, objectivity.

- **Sentiment score:** After the feature score is calculated the polarity of features is displayed and the overall sentiment score is displayed graphically as well as percentile score is generated.

To obtain data from API's we will extract them from open source product review website which provide their API's like reviewing.net

When a product will be searched by the user, we will:

1. Extract product reviews from database(min10-max100)
2. Determine parts of speech using Standard POS Tagger. Split phrases into a vector of sentences and Split sentences into a vector of words.
3. Extract meanings and equivalent score for each word, from scoring module.

Once the scores are generated calculate their overall sentiment score

4. Lastly generate precise statistical representation of the sentiment scores

On the users side he/she will have to log in, once successfully logged the user can search the desired product in the search block and obtain sentiment analysis of that product statistically.

The purpose of this system is to get reviews from the user and rate the products on the basis of their sentiment scores. It will help the users in buying a genuine product and can prevent users from buying fake products.

IV. CONCLUSION

"Customers help customers".

Decisions by the customers are based mostly on the opinion and views of other customer's.On the other side, companies need quick feedback from their customers in order to adapt to their needs in real time. Hence sentiment analysis is required to guide the new buyers to choose the right product which will meet their expectations and also help the manufacturer/ supplier to improve the brand value and thereby increase the demand of its product through millions of satisfied customer's happier sentimental attachment.

V. REFERENCES

[1] "Mining popular menu items of a restaurant from web reviews," in Proceedings of the International Conference onWeb Information Systems andMining (WISM'11) by Y. H. Gu and S. J. Yoo, vol. 6988 of Lecture Notes in Computer Science, pp. 242–250, Springer, 2011.

[2] "Feature and opinion mining for customer review summarization," in Proceedings of the 3rd International Conference on Pattern Recognition and Machine Intelligence (Premi '09), by M. A. Jahiruddin, M. N. Doja, and T. Ahmad, vol. 5909 of Lecture Notes in Computer Science, pp. 219–224, 2009.

[3] S. Huang, X. Liu, X. Peng, and Z. Niu, "Fine-grained product features extraction and categorization in reviews opinion mining," in Proceedings of the 12th IEEE International Conference on Data Mining Workshops (ICDMW'12), pp. 680–686, 2012.

[4] A. Martina, S. Famitha,V.Anithalaskhmi,"Multiple Aspect Ranking using Sentimentclassification"in IJERT, ISSN 2278-0181

[5] www.wikipedia.org